Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Environmental Economics and Management 52 (2006) 469–481
JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT
www.elsevier.com/locate/jeem

Is hypothetical bias universal? Validating contingent valuation responses using a binding public referendum ☆

Robert J. Johnston*

Department of Agricultural and Resource Economics, Connecticut Sea Grant Office, University of Connecticut at Avery Point, 1080 Shennecossett Rd., Groton, CT 06340-6048, USA

Received 19 April 2005
Available online 8 June 2006

Abstract

This study presents a criterion validity test in which stated choice responses are compared to aggregated votes in a subsequent binding referendum. The assessment is characterized by an identical quasi-public good and information content in hypothetical and actual contexts; a genuine and consequential contingent choice survey implemented before the referendum was scheduled or announced; and hypothetical and actual responses representing a large proportion of a well-defined population. The comparison is designed to be simple and unambiguous—no response re-coding is required, no cheap-talk mechanisms are used, and a single choice per respondent parallels the binding referendum. The study is also distinguished by results that show no statistical evidence of hypothetical bias. Findings provide evidence that hypothetical bias is not universal, and suggest potential means to ameliorate hypothetical bias in stated preference research.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Contingent valuation; Contingent choice; Criterion validity; Hypothetical bias; Referendum

1. Introduction

Prior assessments have demonstrated often substantial hypothetical bias in contingent valuation (CV) willingness-to-pay (WTP) estimates, calling into question the appropriateness of such estimates for welfare evaluation. Reviews by Little and Berrens [25], List and Gallett [26], Murphy et al. [34] and Murphy and Stevens [35] illustrate a growing literature devoted to assessments of hypothetical bias. While not all research finds evidence of hypothetical bias [9,22,41,44] and some shows that hypothetical bias may be ameliorated using cheap-talk, certainty adjustments, or other mechanisms [10,12,14,29,36], most research finds significant divergence between stated and actual behavior.[1]

☆This research was funded by the Town of Scituate (RI) Water Study Committee, Connecticut Sea Grant, and the Connecticut Agricultural Experiment Station. The author thanks Patricia Champ, Joshua Duke, Deepak Joglekar, Christopher Modisette and two anonymous reviewers for helpful comments and assistance. Any remaining errors are those of the author. Opinions belong solely to the author and do not imply endorsement by the funding agencies.

*Fax: +860 405 9109.
E-mail address: robert.johnston@uconn.edu.

[1]The meta-analysis of Murphy et al. [34, p. 323], however, does suggest that "hypothetical bias may not be as significant a problem in stated preference analyses as is often thought."

0095-0696/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jeem.2005.12.003

EXHIBIT L

While hypotheses regarding hypothetical bias may be tested using assessments of either convergent or criterion validity,[2] criterion validity tests are generally preferred in cases where appropriate criterion values exist. In CV research, criterion validity tests typically require comparisons between stated behavior or WTP in contingent (hypothetical) contexts and actual behavior in contexts that require some type of real payment [8]. Three general types of criterion validity tests are applicable to public good valuation using CV surveys [44]. These include (1) comparisons of CV responses to behavior in actual markets, (2) comparisons of CV responses to behavior in simulated markets involving actual money transactions, and (3) comparisons of CV responses to voting behavior in binding public referenda.

Notwithstanding the ubiquity of (1) and (2) in the literature, both have characteristics that can lead to ambiguity regarding their implications for criterion validity in public good contexts. Comparisons of CV responses to behavior in actual markets most often involve private goods [e.g., 13, 27]. Hypothetical bias may differ in public and private good contexts, however, leaving the implications of such research for public good valuation unclear [18]. Behavior in simulated markets, in contrast, can diverge from that in comparable real market contexts [44]. Moreover, in cases where simulated markets incorporate voluntary individual or group donation mechanisms [e.g., 10,17,42], results are subject to confounding effects of free-riding and an associated lack of incentive compatibility [9,37]. Notwithstanding the potential capacity of certain payment elicitation mechanisms (e.g., provision point mechanisms) to reduce tendencies to free-ride, such limitations have led some to question whether simulated markets provide appropriate criterion values for hypothetical bias assessments [2,37,44].

Compared to the many private good and simulated market assessments addressing hypothetical bias in CV surveys, only four studies compare hypothetical survey responses to binding referendum results for identical public goods [8,11,44,45].[3] Because these studies compare "intended voting behavior" to referendum results, however, implications for hypothetical bias in genuine[4] stated preference instruments are unclear [39]. For example, as stated by Vossler and Kerkvliet [44, p. 647], "it is still an open question whether our survey responses would have matched the election outcome if all respondents were unaware of an upcoming vote." While Schläpfer et al. [39] compare genuine CV responses and referendum votes, differences between the hypothetical and actual choice contexts in this study lead to questionable comparability of hypothetical and actual behavior.[5] In sum, there are few published assessments of criterion validity that compare hypothetical choices to actual referendum votes, and none that compare genuine CV responses to referendum results for identical goods.

The present study seeks to fill this gap in the literature. It extends work such as that of Vossler and Kerkvliet [44], comparing genuine discrete choice CV responses to aggregated votes in a subsequent, binding public referendum. The assessment is designed to be unambiguous and simple. Hypothetical and actual choice contexts are parallel and consequential, and address the provision of an identical quasi-public good (i.e., the provision of public water to the Village of North Scituate, Rhode Island). Respondents are drawn from the same well-defined population. No re-coding or transformation of survey responses is required, no cheap-talk or certainty adjustments are applied, and a "one vote per survey" format eliminates the need to adjust for correlation or sequence effects among responses.

The study is also distinguished by results that show no statistical evidence of hypothetical bias. While actual and intended voting behavior have often been shown to comport quite closely [e.g., 44], the author is aware of

[2]Convergent validity considers whether a measure (e.g., a contingent valuation WTP estimate) is correlated with other measures of the same theoretical construct. Criterion validity is "concerned with whether the measure ... is related to other measures which may be regarded as criteria" [33, p. 190].

[3]Shabman and Stevenson [40] compare payment card CV responses and voting outcomes. However, the small sample size and open-ended (i.e., not choice based) survey format prevent a reliable comparison of hypothetical and actual voting behavior [44].

[4]Schläpfer et al. [39, pp. 4–5] characterize a *genuine* CV survey as one conducted "before the actual referendum proposition becomes the subject of public debate." They distinguish such surveys from those assessing "intended voting behavior," in which stated preference surveys are conducted after an upcoming referendum has been scheduled and announced. As noted by Schläpfer et al. [39, p. 4], "it appears to be impossible to put respondents in the mindset of a hypothetical choice on the basis of controlled survey information when they are, at the same time, making a real voting choice that is influenced by information from a variety of sources."

[5]The hypothetical context considered a "regional landscape conservation program" funded by a municipal tax, with a "triple-bounded referendum with follow-up" question format. The referendum considered an "increase of cantonal fund for nature and heritage protection" funded by a cantonal tax, with a standard referendum question format [39].

no prior work that shows similar correspondence between binding votes and hypothetical choices in genuine CV contexts. These findings contradict the preponderance of evidence from simulated markets, which suggests the presence of hypothetical bias across a range of CV approaches [35]. Contrary findings from the present study suggest that hypothetical bias is not universal, and hint at potential avenues for the amelioration of such biases in future work.

2. Incentive compatibility and parallelism in comparisons of hypothetical choices and public referenda

For purposes of the present analysis, hypothetical bias is defined as a case in which the proportion or likelihood of 'yes' (or approval) votes associated with a specific and identical good, at a given and identical household cost, differs systematically between a hypothetical choice and actual voting context. Such divergence might be expected in cases where CV survey questions are not incentive compatible, in the sense that truthful preference revelation does not represent an optimal strategy for the survey respondent [7]. However, it has been established that binary (i.e., referendum-type) CV choices are incentive compatible given that certain conditions are met, including the condition that responses are considered by respondents to be consequential [7]. Induced-value laboratory experiments verify this result, showing incentive compatibility in both hypothetical and real referenda, and an equivalent ability to elicit demand information [43].

Beyond induced-value experiments, however, little of the work investigating hypothetical bias conforms unambiguously to conditions for incentive compatibility [7,9]. For example, prior assessments of hypothetical bias often incorporate private goods, voluntary or donation payment mechanisms, non-incentive compatible elicitation formats (e.g., open-ended), hypothetical responses of unclear consequentiality, and other elements associated with non-truthful preference revelation in CV contexts. Such limitations render conclusions regarding hypothetical bias of limited validity [7,9].

While prior comparisons of hypothetical CV choices and binding referendum votes avoid some of these limitations, Champ and Bishop [9, p. 385] note that past attempts to validate CV responses using referendum behavior "have been met with serious limitations," related to a typical lack of parallelism between hypothetical and referendum choice contexts. Similar concerns are detailed by Schläpfer et al. [39]. For example, the presence of an "undecided" choice option [e.g., 8,11,45] can lead to ambiguity regarding the treatment of such responses relative to the 'yes/no' format of public referenda.[6] Other limitations include common differences between information content and presentation in CV surveys and public referenda. For example, where household tax consequences and public good outcomes are rarely specified precisely in public referenda, they are typically specified in detail by CV surveys [39].[7] More generally, typical voters may be less informed than respondents in well-designed CV studies, at least in part due to the confusing language found in many public referenda [9]. The average behavior of voters and survey respondents may also differ due to nonresponse bias or other factors [44], and divergences between goods or payment vehicles may lead to questionable comparability between hypothetical choices and public votes.

Considering these and other potential limitations, two conditions emerge as paramount in the appropriate use of public referenda to validate hypothetical choices. The first is that both hypothetical and actual choices must be viewed as unambiguously consequential, following Carson et al. [7]. The second is that there must be a high degree of consistency between the CV and referendum choice contexts [9].

[6] A reviewer points out that some ballots incorporate "abstain" or "undecided" options, and that the NOAA panel [1] suggests that an "undecided" choice option in a CV survey may be comparable to a public referendum in which a voter can abstain from making a 'yes' or 'no' choice. Vossler et al. [45] also note that some respondents may offer an "undecided" response to a choice question, even when such an option is not explicitly offered by the survey. Considering these factors, the presence of an undecided choice option in a survey may not necessarily represent a barrier to appropriate comparison of hypothetical versus actual responses. It does, however, present the researcher with a choice regarding the treatment of undecided responses [cf. 45].

[7] Of existing comparisons involving identical goods, only Vossler et al. [45] consider a case in which explicit household tax consequences are specified clearly in both the CV survey and public referendum ballot. Vossler and Kerkvliet [44] consider a case in which the projected household cost of the public good was not present on the ballot, but was available in pre-election brochures.

2.1. The North Scituate water supply study

The North Scituate water supply study was designed as a simple, direct comparison of genuine CV survey responses and public referendum votes. Maximal consequentiality and consistency of the CV treatment and referendum were promoted through coordinated development of both the survey and referendum, and the grounding of both in a concrete and specific policy proposal. The study is characterized by: (1) an actual voting choice that may be placed in the same utility maximization context as CV responses, (2) a genuine CV survey implemented before the public referendum was approved or scheduled, (3) a referendum that precisely specifies household tax and quasi-public good outcomes with an elicitation and information content nearly identical to that of the hypothetical survey, (4) consequential hypothetical and actual choice contexts addressing an identical, salient and familiar quasi-public good, and (5) voters and respondents representing a large proportion of a well-defined population.

The study was conducted to assist the Town of Scituate Water Study Committee in assessing public support for provision of public water supply to the Village of North Scituate.[8] The project would have installed and operated infrastructure necessary to provide public water to approximately 450 homes within the Village. Based on State of Rhode Island regulatory requirements, implementation of the proposed water project would have required a majority vote of affected property owners in an officially sanctioned referendum. The Water Study Committee funded the contingent choice survey as a practical means to gauge public WTP for public water in North Scituate, prior to incurring the significant cost required to sanction, promote, schedule, and implement the actual referendum.

Although the survey instrument noted the possibility of a public vote as a possible subsequent step in the process of establishing the water supply project, this was the first indication that any official referendum might be forthcoming. As the survey was designed as a means to assess public preferences—before the official vote was approved or scheduled—it provides a nearly ideal context in which to assess the validity of hypothetical survey responses in a genuine CV context. Appendix A provides the exact text incorporated in the survey cover letter, in which the rationale for the survey was described to respondents.

2.2. The survey instrument

The contingent choice survey was designed in collaboration with the Water Study Committee, and was iteratively revised based on meetings with the Committee and water supply experts, focus groups with local residents, and individual survey pretests. The Committee provided engineering plans for the proposed water project and estimated costs based on consultations with water systems experts. These plans and estimates were used as the basis for both the contingent choice survey and the subsequent binding referendum.

Prior to presenting the contingent choice question, the survey and an accompanying fact sheet provided detailed information concerning public water and the proposed water project. As described to both survey respondents and referendum voters, the water supply project is best characterized as a quasi-public good. The Committee initially envisioned a project in which individual homeowners could "opt-out" of the system—making it more akin to a private good. At the time of the study and subsequently scheduled referendum, however, it was considered a strong possibility that connection to the system would be mandatory [28]. Hence, to reduce the potential for incentive non-compatibility [7,15], the survey presented the choice as involving mandatory payment and water service. In addition, the project would have resulted in other changes characteristic of a public good, including the opening of new areas to potential development (areas unsuitable for private wells), and the nuisance associated with the installation of water supply lines.

The hypothetical choice question asked respondents to vote for or against the proposed water project, given a specified quarterly household water bill that would cover all installation, operation, and maintenance costs. Because costs were not known with 100% certainty, the Committee estimated a realistic range of costs, based

[8]The Village of North Scituate is a geographically well-defined village of approximately 450 homes, located within the larger Town of Scituate.

on data received from neighboring water districts and engineering cost estimates [31].[9] In order to forecast voting behaviors at different cost levels within this range, five different survey versions were produced, each with a different projected cost. The five quarterly cost levels were \$125, \$175, \$250, \$325, and \$425.[10] Respondents were informed that "because costs are not known with 100% certainty, the Water Study Committee has estimated a realistic range of costs. Some receiving this survey will see costs on the high end of this range. Some will see costs on the low end of this range." This statement was included in order to provide a straightforward explanation for the varying cost levels incorporated in versions of the survey received by different households.[11]

After detailing the water supply project and its costs, the dichotomous choice question was worded as follows:

For this final question, please assume that your household's average quarterly bill (paid every 3 months, or four times per year) would be \$(cost level inserted here). Considering the costs and benefits of the proposed public water supply, how would you vote?

- YES—I would vote in favor of the public water supply in my area. I understand that my household would pay approximately \$(cost level inserted here) per quarter in new district water bills.
- NO—I would vote against the public water supply in my area. I understand that my household would continue to pay the costs of operating and maintaining our private well.

3. The data and empirical model

The North Scituate Water Supply Survey was implemented as a mail survey, following the tailored survey design method of Dillman [16]. As only quarterly cost varied across survey versions the experimental design was straightforward, with a roughly equal number of surveys incorporating each of the five quarterly cost levels. Other attributes of the project remained constant over all survey versions. In addition to the contingent choice question, survey responses provided information concerning respondents' satisfaction with the water from their private wells, the ability of their wells to provide an adequate quantity of water year-round, the household's use of water filters and/or bottled water, and other elements of potential relevance.

Surveys were distributed to all homeowners in the region who would be affected by the proposed water supply project. In total, 435 surveys were mailed on July 2, 2001. Of these, 342 surveys were returned, for a final response rate of 79%. Two hundred and ninety-nine surveys were returned based on a preview letter, an initial survey mailing and a reminder letter. An additional 43 were returned subsequent to a final reminder postcard sent on July 28.[12] This distinction is noted as it appears that there is a systematic difference between responses received before and after the reminder postcard. Of all returned surveys, 328 include a completed response to the contingent choice question, and 304 provide sufficient response data for all variables to enable inclusion in the empirical model of CV (hypothetical) responses.

[9]Costs were estimated assuming that all installation and capital costs would be funded through a 30-year bond issued by the Town of Scituate (5–7% interest rate). Quarterly bond payments per household were added to an estimated range of quarterly (per household) operating costs from neighboring water districts, to obtain a range of total costs. A quarterly cost of \$250 was the mean value of this range, rounded to the nearest \$10.

[10]While the survey presented costs on a quarterly basis, the referendum presented costs on an annual basis. Quarterly costs were specified in the survey to match the frequency with which water bills would be paid by households. The median quarterly cost included in the survey (\$250/quarter) was designed to match the expected annual cost (\$1000/year) of the project.

[11]A reviewer notes that such statements are not typical in CV surveys, and could lead to cases in which a survey respondent might anchor or respond to a cost level distinct from that presented in the survey, if he or she perceives the presented level to be unrealistic. This possibility notwithstanding, focus groups and pretests showed no evidence of such behavior.

[12]Observable characteristics (e.g., household size, supply characteristics of private wells, use of bottled water, etc.) of respondents returning surveys before and after the reminder postcard are statistically indistinguishable.

Case 4:05-cv-00329-GKF-PJC Document 2272-13 Filed in USDC ND/OK on 06/19/2009 Page 6 of 13

3.1. The North Scituate water supply referendum

Subsequent to the implementation of the survey, a binding referendum to authorize the water supply project was sanctioned and scheduled by the Town of Scituate. The referendum concerned "whether to establish an independent water district...[with an]...approximate cost per year to the people receiving the service [of] $1000," an annual cost matching the median $250/quarter cost presented to CV survey respondents. The referendum was held during a Special Financial Town Meeting on November 13, 2001 [28].[13] The water supply project was identical to that considered in the hypothetical survey instrument implemented approximately 4 months prior. Description of the project in the referendum was drawn from the same information found in the hypothetical survey instrument. All those eligible to vote in the Town Meeting had already received the prior survey describing the water supply project.

Aggregate data for the November 13, 2001 referendum were obtained from official minutes of the Special Financial Town Meeting [28]. After the provision of a verbal description of the project and a limited question-and-answer session,[14] the vote was initiated at 8:45 p.m. Final ballot results, at an anticipated household cost of $250 per quarter, were 122 in favor and 145 opposed, representing 45.69% approval.

3.2. The empirical model

The data are analyzed using a standard random utility model [19,32]. The model presumes that respondents choose whether to vote for mandatory public water based on the difference in utility, dU, resulting from public water supply and that resulting from retaining private (well) supply, given by

$$dU = U_p(\boldsymbol{W}_p, Y - C_p, \boldsymbol{D}) - U_w(\boldsymbol{W}_w, Y - C_w, \boldsymbol{D}) = v_p(\boldsymbol{W}_p, Y - C_p, \boldsymbol{D}) - v_w(\boldsymbol{W}_w, Y - C_w, \boldsymbol{D}) - [\varepsilon_w - \varepsilon_p] = dv - \theta. \quad (1)$$

Here, $U_p(\cdot)$ represents utility realized from the provision of public water and $U_w(\cdot)$ represents utility realized from the retention of private water. The functions $v_p(\cdot)$ and $v_w(\cdot)$ represent the empirically measurable components of utility associated with public and well water supply, respectively, while ε_p and ε_w represent the associated stochastic or unobservable components.

Utility from public water supply is assumed to be determined by the attributes of public water (vector $\boldsymbol{W}_p$, assumed fixed across households), the household's demographic attributes (vector $\boldsymbol{D}$), and household income (Y) minus the cost of public water (C_p). Utility from private well water is determined by the attributes of the household's well water supply (vector $\boldsymbol{W}_w$, which may vary across households), demographic attributes, and household income minus the cost of operating a private well (C_w). Based on interviews with the Water Study Committee and focus groups, this cost (C_w) is assumed to be fixed and constant across households.

Given a standard linear approximation for $v_p(\cdot)$ and $v_w(\cdot)$, and assuming that $\boldsymbol{W}_p$ and C_w are fixed and constant across all observations, dv is given by the reduced-form specification

$$dv = \beta_0 + \boldsymbol{\beta}_1(\boldsymbol{W}_w) + \beta_2(C_p) + \boldsymbol{\beta}_3(\boldsymbol{D}), \quad (2)$$

where the β_i represent parameters (or vectors of parameters) to be estimated. Assuming that θ in (1) follows a logistic distribution, one may estimate parameters in (2) using a standard logit model [30]. As each survey includes a single choice question, observations (i.e., survey responses) may be treated as independently and identically distributed (*iid*). Table 1 characterizes independent variables included the statistical model.

[13] It was emphasized that as the project proceeded, certain additional elements of the proposed water supply project would likely be placed "before the voters of the district again for final approval" [28]. The specific elements that would require additional voter approval were left unspecified. However, it was made clear that prior to the completion of the project (a multiple year task), subsequent votes would likely be necessary to address specific details of the project. Regarding the issue of mandatory versus voluntary participation in the public water system, once established, it was emphasized that mandatory connection with the system "was a possibility."

[14] Five individuals posed questions at the meeting. Most addressed minor procedural issues such as the residency status of those serving on the Water Study Committee, the procedure for absentee ballots, and whether those serving on the Committee would be paid or volunteer. Also asked were questions for which no answers could be provided at the time of the meeting, including the potential location of water tanks, and whether liens could be placed on the properties of homeowners who refused to connect [28]. No new information of a significant nature was revealed during the session, beyond that already provided in the survey instrument.

EXHIBIT L

Table 1
Model variables and summary statistics

Variable	Definition	Mean	Std. dev.
Bottled	Binary variable indicating that the household uses bottled water for drinking (1 = yes, 0 = no).	0.140	0.348
Supply	Binary variable identifying households whose well provides an adequate quantity of water during all months of the year (1 = yes, 0 = no).	0.879	0.327
H_size	Number of persons living in the household.	2.835	1.462
Late	Binary variable identifying surveys received "late," or after a reminder postcard was mailed on July 28, 2001.	0.126	0.332
Cost	Estimated quarterly household cost of the proposed water supply project, including all installation, maintenance, and operation costs.	248.099	101.960
Cost × late	Interaction between *cost* and *late*.	32.749	95.415

3.3. Comparability of respondent and voter samples

Identities of survey respondents are known; individual identities of referendum voters are not available.[15] Although the representation of households in the voter sample is unknown, the 267 voters likely represent between 31% and 61% of the 435 affected households, depending on assumptions regarding the number of eligible voters per average household in North Scituate.[16] This compares to 79% of households represented by the survey sample. The lack of experimental control over individuals in the referendum sample leaves open the possibility, even if remote, that findings presented below are the result of differences between the two samples. Nonetheless, even in the worst-case scenario, both the survey (79%) and the referendum (31–61%) represent a considerable proportion of affected households. Assuming that incentives for participation in the vote and participation in the survey are similar, it seems reasonably likely—although not certain—that the underlying preferences of the survey and voter samples are comparable.

4. Model results

Data identifying the referendum votes of specific individuals are not available, preventing a detailed assessment of individual behavior such as that of Vossler and Kerkvliet [44]. Hence, to assess criterion validity, CV responses are compared to aggregate referendum results. Table 2 illustrates the raw choice results associated with each of the five quarterly cost levels. The data are well-behaved, with the proportion of 'yes' responses decreasing with increases in cost. The data reveal that 48.44% of survey respondents shown a cost of $250/quarter voted in favor of the water supply project. This closely approximates the 45.69% of voters in the referendum who also voted in favor of the project, at the same quarterly cost. A test of equal proportions (Table 2) fails to reject the equality of raw hypothetical and actual vote proportions at a $250/quarter cost ($p = 0.69$). While not representing a full, formal assessment of hypothetical bias, the statistical equivalence of raw voting outcomes suggests a high degree of association between hypothetical and real choices.[17]

[15]Despite numerous attempts to obtain the identification of referendum voters from local and state sources, we were informed that the data had been discarded upon official certification of the vote. In-person exploration of Town of Scituate records revealed the same result—data showing the identities of referendum voters had not been archived.

[16]If one assumes exactly one eligible voter per household (a likely lower bound), then the 267 votes represents 61% of the 435 households. If one assumes two eligible voters per household (a likely upper bound), then 267 votes would represent a *minimum* of 31% of the 435 households.

[17]Statistical tests presented in this and subsequent sections presume independence between the samples of survey responses and actual votes. This is unavoidable because only aggregate voting results are available and hence the correlation between survey responses and votes is unknown. If positive correlation exists (because of overlap between respondents in the two samples), the reported confidence bounds will be overly narrow, increasing the likelihood of rejecting the null hypothesis of zero hypothetical bias. Because we fail to reject the null hypothesis in all cases nonetheless, the failure to address potential (positive) correlation between actual and hypothetical votes cannot alter our fundamental hypothesis test results.

Table 2
Raw results: proportion of 'yes' votes at different quarterly cost levels

	Proportion of 'Yes' votes		Test of equal proportions (H_0: Survey proportion = Referendum proportion)
	Hypothetical survey responses	Binding referendum responses	
Quarterly cost level[a]			
\$125 ($n = 72$)	0.653	—	
\$175 ($n = 75$)	0.587	—	
\$250 ($n = 64$ survey; $n = 267$ referendum)	0.484	0.457	$Z = 0.396, p = 0.692$
\$325 ($n = 71$)	0.409	—	
\$425 ($n = 46$)	0.326	—	

[a]Numbers in parentheses indicate the total number of surveys received at each stated cost level.

Table 3
Maximum likelihood logit results: contingent choice model

Variable	Coefficient	Std. err.	Prob > | z |
Intercept	5.044	0.872	<0.001
Bottled	0.894	0.414	0.031
Supply	−3.531	0.675	<0.001
H_size	0.174	0.093	0.062
Late	−2.309	0.932	0.013
Cost	−0.011	0.002	<0.001
Late × cost	0.012	0.004	0.001
-2 Ln L χ^2: model (df = 6)	87.578		<0.001
Pseudo R^2	0.208		
N (obs.)	304		

Results of the logit model are presented in Table 3. The model is statistically significant at $p<0.001$, with 77% of observed response pairs concordant with model predictions. All independent variables are statistically significant at standard levels. A restricted specification was also estimated, from which two variables (*late*; *late* × *cost*) were excluded.[18] These variables allow estimated parameters for the equation intercept and program cost to vary systematically according to whether the survey was received after the reminder postcard was mailed on July 28. A likelihood ratio test of restrictions in the restricted model rejects the null hypothesis of zero joint influence ($\chi^2 = 12.32$, $p = 0.002$), indicating that there are systematic differences between the preferences of those who returned surveys before and after July 28. Model results also suggest that late respondents were less influenced by the household cost of the project, as indicated by the near-zero sum of coefficients on *cost* and *late* × *cost*. While the explanation for this finding is unknown, it suggests that the timing of survey responses may influence the potential for hypothetical bias, as in this case it influences the responsiveness of choices to program cost. Given statistical significance of restrictions noted above, the unrestricted model is used for assessments of criterion validity.

4.1. Assessments of criterion validity and hypothetical bias

As noted above, referendum results are characterized by a single point: 45.69% approval of the proposed project at an estimated household cost of \$250 per quarter. This limits the extent of comparisons that may be

[18]Restricted model results are suppressed for the sake of parsimony, but are available from the author upon request.

Table 4
Assessment of hypothetical bias: empirical results

	Mean prediction[a]	On-time prediction[a]
Proportions of 'Yes' votes at $250/quarter cost[b]		
Contingent choice—hypothetical	0.497	0.477
	(0.425, 0.565)	(0.402, 0.550)
Referendum—actual	0.457	0.457
	(0.397, 0.517)	(0.397, 0.517)
Point estimate difference	0.041	0.020

Statistical test of hypothetical vs. actual proportions ($250/quarter cost)

	Test statistic	*p*-Value	Test statistic	*p*-Value
t-Test for for H_0: [hypothetical = referendum proportion]	$t = 0.862$	0.389	$t = 0.416$	0.677

[a]The "mean" prediction is generated from unrestricted logit model results (Table 3), assuming mean values for all model variables except *cost*. The "on-time" prediction is generated from the same model results assuming that *late* = 0. Results represent the proportion of representative respondents predicted to support the water supply project at a cost of $250/quarter.
[b]Numbers in parentheses are 95% confidence intervals.

made between hypothetical and referendum responses, as a complete preference function may only be estimated for hypothetical responses. Nonetheless, even within these limitations, one may assess differences between responses in the two choice contexts.

The most direct comparison involves the probability of supporting the proposed water supply project (or, alternatively, the proportion of representative respondents who would support the project). The proportion of votes supporting the project is known in the referendum—0.457 at $250/quarter. Given an identical cost and mean values for other model variables, the comparable proportion of support predicted by the CV model is 0.497. If only on-time survey responses are considered (i.e., *late* = 0), the proportion falls to 0.477. Subsequent discussion refers to these as the "mean" and "on-time" predictions of the CV model, respectively.

Confidence intervals and standard errors (Table 4) for the predicted hypothetical proportions of support at $250/quarter (mean and on-time) are calculated following Krinsky and Robb [24]. The procedure randomly draws 10,000 sets of coefficient estimates from the estimated maximum likelihood parameter distributions. The predicted proportion of project support (at $250/quarter) is calculated for each draw, resulting in an empirical distribution from which confidence intervals and standard errors are derived. Resulting standard errors are 0.036 for the mean prediction and 0.038 for the on-time prediction. Anderson–Darling tests fail to reject the null hypothesis of normality for the both distributions, suggesting that standard hypothesis tests are appropriate [38].[19] For the proportion of actual support, 95% confidence intervals are estimated based on the standard error (0.030) calculated directly from referendum data, following established methods [21].

Hypothetical and actual proportions are compared using *t*-tests [cf. 44]. Based on standard errors estimated above, we fail to reject the null hypothesis of equality between the mean hypothetical and actual proportions of 'yes' votes at $250/quarter ($p = 0.39$ and 0.68 for the mean and on-time predictions). That is, hypothesis test results indicate that there is no statistically significant difference between CV responses and referendum votes (Table 4).[20]

[19]Anderson–Darling tests fail to reject the null hypothesis of normality for the resampling distribution of proportions for both the mean model prediction ($p > 0.25$) and the case in which *late* = 0 ($p > 0.25$). Parallel results are generated by other common tests of normality (e.g., Kolmogorov–Smirnov, Cramer–von Mises).
[20]Wald tests generate equivalent results, failing to reject the null hypothesis that the predicted hypothetical proportion of supporting votes at $250/quarter is equal to 0.4569.

5. Implications and discussion

This is the first published assessment in which hypothetical responses in a genuine CV context have been compared to referendum results for identical programs. Results of this comparison show no evidence of statistically significant hypothetical bias. The high degree of parallelism between the hypothetical and referendum contexts suggests that model results provide a valid, if limited assessment of criterion validity. This result begs the question as to why hypothetical bias is not present, despite clear evidence of hypothetical bias in simulated market studies.

Although the present data do not allow unambiguous determination of the reasons why hypothetical bias is not significant, there are a variety of potential explanations that correspond to prior findings in the literature. These include: (1) the salience and familiarity of the good and equivalence of information content in hypothetical and actual choice contexts and (2) the explicit linkage between the survey and an official government process.

5.1. *Salience and familiarity of the good and equivalence of information content*

According to Arrow et al. [1, p. 4605], "[i]f CV surveys are to elicit useful information about WTP, respondents must understand exactly what it is they are being asked to value." The survey and focus group literature suggests that first hand experience with natural resources may influence values and survey responses [6,23], and that familiarity may lead to closer correspondence between intentions and behavior [5, p. 128; 33, p. 186]. Here, the survey presents a quasi-public good that is clearly familiar and salient. Many respondents will have had direct experience with the attributes of this good—public water supply—in other homes or businesses. This stands in contrast to goods often valued using CV methods, which may be unfamiliar and with which respondents may have little direct experience.

It has also been established that information provision can influence WTP estimates [3,4,20] and that information in CV surveys often differs from that in public referenda [9,39]. Here, the information content underlying hypothetical and actual choices is nearly identical. All households in the affected neighborhood were provided with extensive information regarding the proposed water supply project as part of the survey, and information provided during the public meeting was drawn from informational materials developed for the survey.[21] The extensive and equivalent information provided in both choice contexts—combined with the familiarity and salience of the good—may have contributed to the close correspondence between hypothetical and actual choices observed in the present case.

5.2. *Explicit linkage between the survey and an official government process*

The survey described here is unusual in that it was implemented as part of an official government consideration of a specific public project. Also, while the decision to hold the referendum had not been made at the time of survey implementation, the possibility of a forthcoming referendum was mentioned as part of the survey materials (see Appendix A). Hence, while the survey is appropriately characterized as a genuine CV instrument, the association with a government process and the possibility of a subsequent referendum may have placed respondents in a mindset closer to that which would apply during a binding vote.

If the above supposition is indeed valid, the simple mention that a survey has been commissioned by a government agency—or that a future referendum is possibility—might be sufficient to place respondents in a mindset similar to that which would occur during an analogous binding vote. While survey results cannot confirm such a hypothesis, it is consistent with both the results found here and the correspondence between referendum votes and survey responses in cases where the referendum has been announced prior to survey implementation [44]. Of course, such statements are only appropriate in cases where they are indeed factual,

[21] It is of course possible that additional information may have been available as a result of media reports that surfaced subsequent to the survey and before the referendum. However, much of the information provided in media reports was drawn from the same planning documents upon which the survey information materials were based.

rendering this suggestion mute for CV surveys not associated with actual policy proposals or potential future referenda.

This reasoning is also consistent with Carson et al.'s [7] arguments concerning the role of consequentiality in CV surveys, defined as a case in which "survey results are seen by the agent as potentially influencing agency action" and in which the agent treats the survey as "an opportunity to influence those actions." Here, survey materials present a clear description of the role of survey results in the policymaking process. Moreover, the agency in consideration was a local government body comprised of known and largely trusted individuals. Hence, respondents may have viewed the survey as being more highly consequential than typical CV instruments. Based on the arguments of Carson et al. [7], this may have led to a greater incentive for respondents to answer truthfully concerning their willingness to support the water supply project.

6. Conclusion

While reasons for the lack of hypothetical bias cannot be unambiguously established in the present case, this study suggests a number of possibilities for the amelioration of such biases in broader CV research. Among the possibilities are that reductions in hypothetical bias may be related to the salience and familiarity of the good in question, the equivalence of information content in hypothetical and binding choice contexts, and an explicit linkage between the survey and an official government process. Additional research is required, however, to establish which, if any of these factors might provide a practical means to reduce or eliminate hypothetical biases in broader CV research. In contrast, subsequent research might also find that present results are due primarily to unique elements of the present study, and that these results hold little promise for improvements to CV methods in general.

Additional findings of this analysis include the apparent relationship between survey response dates and the relative weight given to the payment vehicle. While the cause of this unexpected relationship is unknown, it does suggest the possible existence of hitherto unexplored correlations between survey response behaviors and hypothetical bias. It is hoped that future research may assess whether similar patterns hold in broader CV analysis, and clarify potential implications of such findings for the criterion validity of CV responses.

Empirical results shown here may offer more in terms of questions and possibilities than unambiguous proof that hypothetical biases may be overcome in broader CV research. While results suggest that hypothetical bias can be avoided in some instances, there remains a possibility that results reflect other influences such as divergent rates of nonresponse or other unidentified differences between observed voters and survey respondents, or heretofore unidentified divergences between the hypothetical and actual choice contexts. Such possibilities notwithstanding, model results suggest that researchers should neither abandon the search for means to ameliorate hypothetical biases, nor should accept such biases as a foregone conclusion.

Appendix A. Text of selected survey materials

The following appendix provides the text incorporated in the survey cover letter, in which the rationale for the survey was described to respondents.

A.1. Text of the cover letter

"I am writing to ask for your input on an important decision that affects your household—the potential creation of a water district to provide public water to the Village of North Scituate and points North along Route 116. The Town of Scituate Water Study Committee has prepared this survey to assess public support or opposition to this project. You have received this survey because your home is in the region under consideration for public water.

The proposed water district will not proceed without the majority support of property owners who live within the boundaries of the proposed district. This survey is the first step in determining whether such support exists. Depending on survey results, the next step would be a public information forum, followed by an official vote of affected homeowners in November, 2001. Only homeowners within the proposed water district would be allowed to vote for or against the project.

Included in this package are a one-page survey and a brief fact sheet providing some important information regarding the proposed water district. It is important that we hear from all residents who may be affected by the proposed water supply project. Your answers to this survey are strictly confidential and anonymous. If you have any questions or concerns regarding the survey, please do not hesitate to contact The Town of Scituate Water Study Board at (xxx) xxx-xxxx. We hope that you will take the time to complete this important survey."

References

[1] K. Arrow, R. Solow, E. Leamer, P. Portney, R. Rander, H. Schuman, Report of the NOAA panel on contingent valuation, Fed. Reg. 58 (1993) 4602–4614.

[2] I.J. Bateman, R.T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Leee, G. Loomes, S. Mourato, E. Ozdemiroglu, D.W. Pierce, R. Sugden, J. Swanson, Economic Valuation with Stated Preference Surveys: A Manual, Edward Elgar, Northampton, MA, 2002.

[3] J.C. Bergstrom, J.R. Stoll, Application of experimental economics concepts and precepts to CVM field survey procedures, Western J. Agric. Econom. 14 (1) (1989) 98–109.

[4] J.C. Bergstrom, J.R. Stoll, A. Randall, Information effects in contingent markets, Am. J. Agric. Econom. 71 (3) (1989) 321–323.

[5] R. Blamey, J. Bennett, J.J. Louviere, M. Morrison, Green product choice, in: J. Bennett, R. Blamey (Eds.), The Choice Modelling Approach to Environmental Valuation, Edward Elgar, Northampton, MA, 2001.

[6] T.A. Cameron, J. Englin, Respondent experience and contingent valuation of environmental goods, J. Environ. Econom. Manage. 33 (3) (1997) 296–313.

[7] R.T. Carson, T. Groves, M.J. Machina, Incentive and Informational Properties of Preference Questions, Working Paper, Department of Economics, University of California, San Diego, 2000.

[8] R.T. Carson, W.M. Hanemann, R.C. Mitchell, The Use of Simulated Political Markets to Value Public Goods, Working Paper, Economics Department, University of California, San Diego, 1986.

[9] P.A. Champ, R.C. Bishop, Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias, Environ. Resour. Econom. 19 (4) (2001) 383–402.

[10] P.A. Champ, R.C. Bishop, T.C. Brown, D.W. McCollum, Using donation mechanisms to value nonuse benefits from public goods, J. Environ. Econom. Manage. 33 (2) (1997) 151–162.

[11] P.A. Champ, T.C. Brown, A comparison of contingent and actual voting behavior, in: Proceedings from the W-133 Benefits and Cost Transfer in Natural Resource Planning, 10th Interim Report, Rocky Mountain Research Station, USDA Forest Service, 1997.

[12] P.A. Champ, R. Moore, R.C. Bishop, Hypothetical bias: the mitigating effects of certainty questions and cheap talk, Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, 2004.

[13] R.G. Cummings, G.W. Harrison, E. Rutström, Homegrown values and hypothetical surveys: do dichotomous choice questions elicit real economic commitments?, Am. Econom. Rev. 85 (1) (1995) 260–266.

[14] R.G. Cummings, L.O. Taylor, Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method, Am. Econom. Rev. 89 (3) (1999) 649–665.

[15] R.G. Cummings, S. Elliott, G.W. Harrison, J. Murphy, Are hypothetical referenda incentive compatible?, J. Political Economy 105 (3) (1997) 609–621.

[16] D.A. Dillman, Mail and Internet Surveys: The Tailored Design Method, Wiley, New York, NY, 2000.

[17] V. Foster, I.J. Bateman, D. Harley, Real and hypothetical willingness to pay for environmental preservation: a non-experimental comparison, J. Agric. Econom. 48 (2) (1997) 123–138.

[18] J.A. Fox, J.F. Shogren, D.J. Hayes, J.B. Kliebenstein, CVM-X: calibrating contingent values with experimental auction markets, Am. J. Agric. Econom. 80 (3) (1998) 455–465.

[19] W.M. Hanemann, Welfare evaluations in contingent valuation experiments with discrete responses, Am. J. Agric. Econom. 66 (3) (1984) 332–341.

[20] J.P. Hoehn, A. Randall, The effect of resource quality information on resource injury perceptions and contingent values, Resour. Energy Econom. 24 (ER1-2) (2002) 13–31.

[21] R.L. Iman, W.J. Conover, Modern Business Statistics, second ed., Wiley, New York, 1989.

[22] M. Johannesson, Some further experimental results on hypothetical versus real willingness to pay, Appl. Econom. Lett. 4 (1997) 535–536.

[23] R.J. Johnston, T.F. Weaver, L.A. Smith, S.K. Swallow, Contingent valuation focus groups: insights from ethnographic interview techniques, Agric. Resour. Econom. Rev. 24 (1) (1995) 56–69.

[24] I. Krinsky, A.L. Robb, On approximating the statistical properties of elasticities, Rev. Econom. Stat. 68 (4) (1986) 715–719.

[25] J. Little, R. Berrens, Explaining disparities between actual and hypothetical stated values: further investigation using meta-analysis, Econom. Bull. 3 (1) (2004) 1–13.

[26] J.A. List, C. Gallet, What experimental protocols influence disparities between actual and hypothetical stated values?, Environ. Resour. Econom. 20 (3) (2001) 241–254.

[27] J.A. List, J.F. Shogren, Calibration of the difference between actual and hypothetical valuations in a field experiment, J. Econom. Behav. Organiz. 37 (2) (1998) 193–206.

EXHIBIT L

[28] M.M. Long, Minutes of Special Financial Town Meeting, Office of the Town Clerk, Town of Scituate, Rhode Island, 2001 (November 13).
[29] J. Loomis, T. Brown, B. Lucero, G. Peterson, Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay, Land Econom. 72 (4) (1996) 450–461.
[30] G.S. Maddala, Limited-Dependent and Qualitative Variables in Econometrics, Econometric Society Monographs No. 3, Cambridge University Press, Cambridge, 1983.
[31] Maguire Group Inc., Report on Water Supply for Scituate, Rhode Island, 1975 Revised to 2000, Final Report June 30, 225 Chapman Street, Providence, RI 02905, 2000.
[32] K.E. McConnell, Models for referendum data: the structure of discrete choice models for contingent valuation, J. Environ. Econom. Manage. 18 (1) (1990) 19–34.
[33] R.C. Mitchell, R.T. Carson, Using Surveys to Value Public Goods: The Contingent Valuation, Method, Resources for the Future, Johns Hopkins University Press, Baltimore, 1989.
[34] J.J. Murphy, P.G. Allen, T.H. Stevens, D. Weatherhead, A meta-analysis of hypothetical bias in stated preference valuation, Environ. Resour. Econom. 30 (3) (2005) 313–325.
[35] J.J. Murphy, T.H. Stevens, Contingent valuation, hypothetical bias, and experimental economics, Agric. Resour. Econom. Rev. 33 (2) (2004) 182–192.
[36] J.J. Murphy, T.H. Stevens, D. Weatherhead, Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism?, Environ. Resour. Econom. 30 (3) (2005) 327–343.
[37] G.L. Poe, J.E. Clark, D. Rondeau, W.D. Schulze, Provision point mechanism and field validity tests of contingent valuation, Environ. Resour. Econom. 23 (1) (2002) 105–131.
[38] G.L. Poe, K.L. Giraud, J.B. Loomis, Computational methods for measuring the difference of empirical distributions, Am. J. Agric. Econom. 87 (2) (2005) 353–364.
[39] F. Schläpfer, A. Roschewitz, N. Hanley, Validation of stated preferences for public goods: a comparison of contingent valuation survey response and voting behavior, Ecol. Econom. 51 (1) (2004) 1–16.
[40] L. Shabman, K. Stephenson, Searching for the correct benefit estimate: empirical evidence for an alternative perspective, Land Econom. 72 (4) (1996) 433–449.
[41] V.K. Smith, C. Mansfield, Buying time: real and hypothetical offers, J. Environ. Econom. Manage. 36 (3) (1998) 209–224.
[42] M.A. Spencer, S.K. Swallow, C.J. Miller, Valuing water quality monitoring: a contingent valuation experiment involving hypothetical and real payments, Agric. Resour. Econom. Rev. 27 (1) (1998) 28–42.
[43] L.O. Taylor, M. McKee, S.K. Laury, R.G. Cummings, Induced-value tests of the referendum voting mechanism, Econom. Lett. 71 (1) (2001) 61–65.
[44] C.A. Vossler, J. Kerkvliet, A criterion validity test of the contingent valuation method: comparing hypothetical and actual voting behavior for a public referendum, J. Environ. Econom. Manage. 45 (3) (2003) 631–649.
[45] C.A. Vossler, J. Kerkvliet, S. Polasky, O. Gainutdinova, Externally validating contingent valuation: an open-space survey and referendum in Corvallis Oregon, J. Econom. Behav. Organiz. 51 (2) (2003) 261–277.